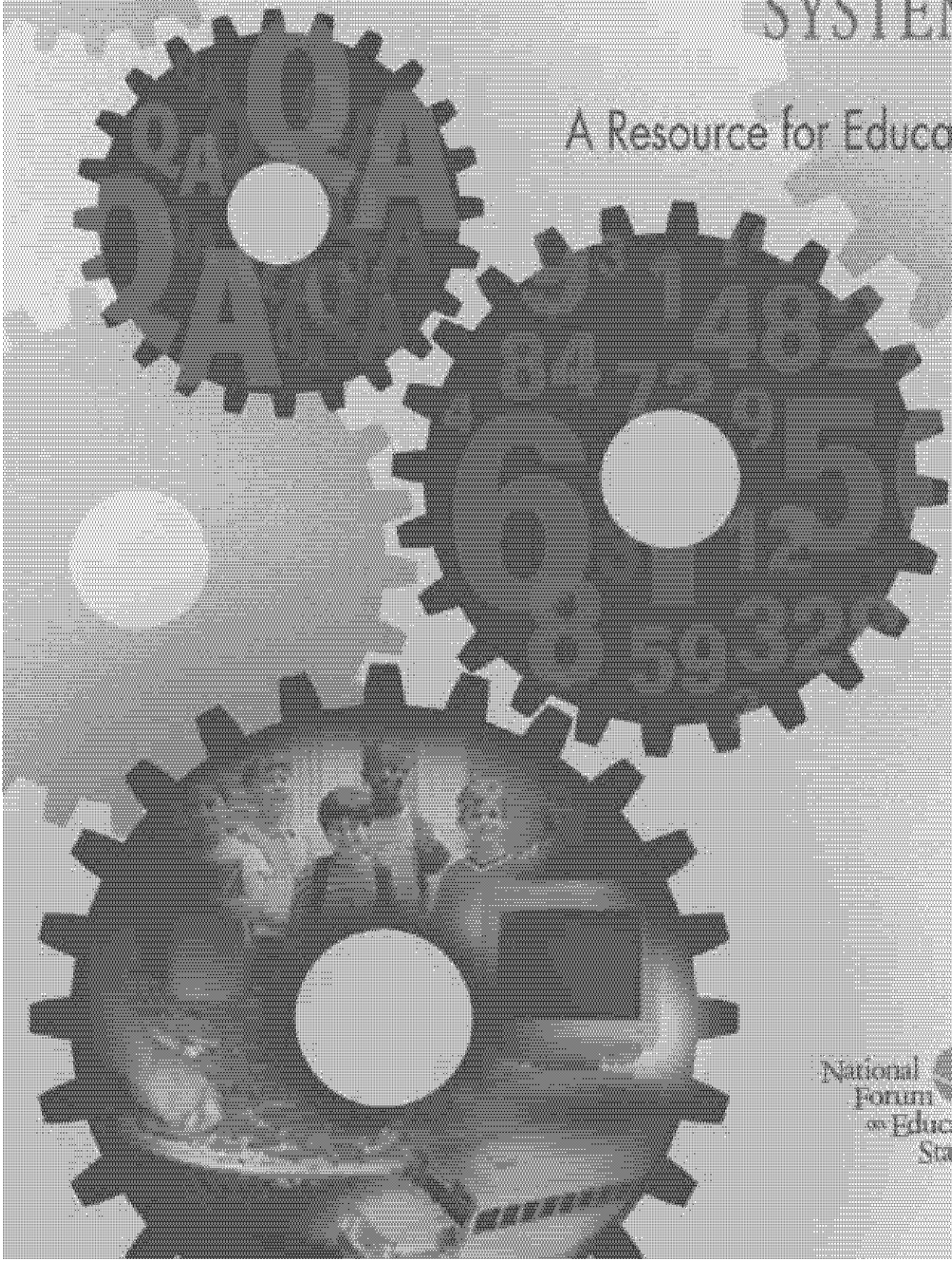
Forum Guide to DECISION SUPPORT SYSTEMS

A Resource for Educators



National Forum on Education Statistics

NATIONAL COOPERATIVE EDUCATION STATISTICS SYSTEM

The National Center for Education Statistics established the National Cooperative Education Statistics System (Cooperative System) to assist in producing and maintaining comparable and uniform information and data on early childhood education and elementary and secondary education. These data are intended to be useful for policymaking at the federal, state, and local levels.

The National Forum on Education Statistics, among other activities, proposes principles of good practice to assist state and local education agencies in meeting this purpose. The Cooperative System and the National Forum on Education Statistics are supported in these endeavors by resources from the National Center for Education Statistics.

Publications of the National Forum on Education Statistics do not undergo the formal review required for products of the National Center for Education Statistics. The information and opinions published here are the product of the National Forum on Education Statistics and do not necessarily represent the policy or views of the U.S. Department of Education or the National Center for Education Statistics.

September 2006

This publication and other publications of the National Forum on Education Statistics may be found at the National Center for Education Statistics website.

The NCES World Wide Web Home Page is *http://nces.ed.gov*
The NCES World Wide Web Electronic Catalog is *http://nces.ed.gov/pubsearch*
The Forum World Wide Web Home Page is *http://nces.ed.gov/forum*

Suggested Citation:
National Forum on Education Statistics. (2006). *Forum Guide to Decision Support Systems: A Resource for Educators* (NFES 2006–807). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

For ordering information on this report, write:
U.S. Department of Education
ED Pubs
P.O. Box 1398
Jessup, MD 20794-1398

Or call toll free at 1-877-4ED-PUBS or order online at *http://www.edpubs.org.*

Technical Contact:
Ghedam Bairu
202-502–7304
ghedam.bairu@ed.gov

Task Force Members

This document was developed through the National Cooperative Education Statistics System, funded by the National Center for Education Statistics (NCES) of the U.S. Department of Education, and produced by a volunteer task force of the National Forum on Education Statistics (an entity of the National Cooperative Education Statistics System). A list of task force members follows.

Co-Chairs

Tom Ogle
Missouri Department of
Elementary and Secondary Education
Jefferson City, Missouri

Raymond Yeagley
Northwest Evaluation Association
Lake Oswego, Oregon

Members

Bethann Canada
Virginia Department of Education
Richmond, Virginia

Bertha Doar
Rockwood School District
Eureka, Missouri

Patricia Eiland
Alabama State Department of Education
Montgomery, Alabama

Thomas Purwin
Jersey City Public Schools
Jersey City, New Jersey

Lee Rabbitt
North Kingstown School Department
North Kingstown, Rhode Island

Jeff Stowe
Arizona Department of Education
Phoenix, Arizona

Project Officer

Ghedam Bairu
National Center for Education Statistics

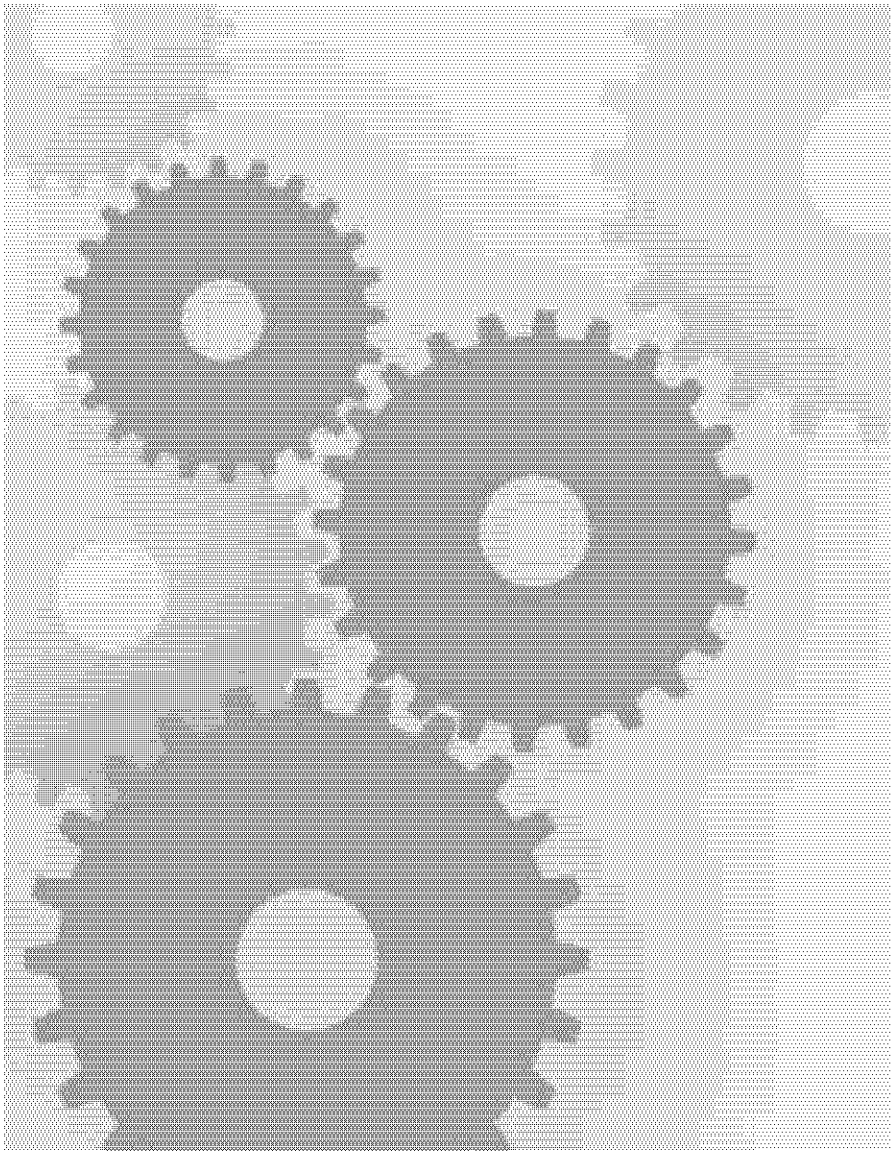
Acknowledgments

Many local, state, and federal education agency officials contributed to the development of this guide. The members of the Forum's Decision Support System Literacy Task Force wish to acknowledge this support. The task force would also like to thank the Technology (TECH) Committee and all Forum members for their help in developing this document.

A web resource published by Daniel J. Power strongly influenced the task force's early work. Carol Dodd (Los Angeles Unified School District) provided suggestions about the use of decision support systems to improve communications between administrators and teachers. Larry Fruth (Schools Interoperability Framework Association) shared expertise during discussions about interoperability. Tim Magner (U.S. Department of Education, Office of Educational Technology) provided expert technical feedback that considerably improved this guide. Bob Bellamy (Better School Business, LLC) also provided technical expertise. Andy Rogers (Quality Information Partners) and Stephanie Rovito (Education Statistics Services Institute) prepared the first draft of this guide. Tom Szuba (Quality Information Partners) offered useful recommendations to improve the document, and helped write it. The task force is also grateful to Lee Hoffman and Emmanuel Sikali (NCES) for reviewing the document. Frances Erlebacher edited the document and the Creative Shop provided layout and design services.

Throughout this document, there are references to several publications produced by the National Forum on Education Statistics. The task force highly recommends these documents as they represent the work of education practitioners from across the country. They can be accessed at *http://nces.ed.gov/forum/publications.asp*.

Table of Contents



Foreword

The Decision Support System Literacy Task Force of the National Forum on Education Statistics (Forum) is pleased to introduce the *Forum Guide to Decision Support Systems: A Resource for Educators.* This document was developed by educators for educators to remedy the lack of reliable, objective information available to the education community about decision support systems. The authors hope it will help readers better understand what decision support systems are, how they are configured, how they operate, and how they might be implemented in an education institution.

Like other Forum guides, this document was prepared by Forum members–representatives of local and state education agencies, federal agencies, and national organizations involved in education data collection and use. This work was supported by the National Center for Education Statistics of the U.S. Department of Education.

The National Forum on Education Statistics provides an arena for local, state, and national leaders in the education data community to discuss issues, address problems, develop resources, and consider new approaches to improving data collection and utility.

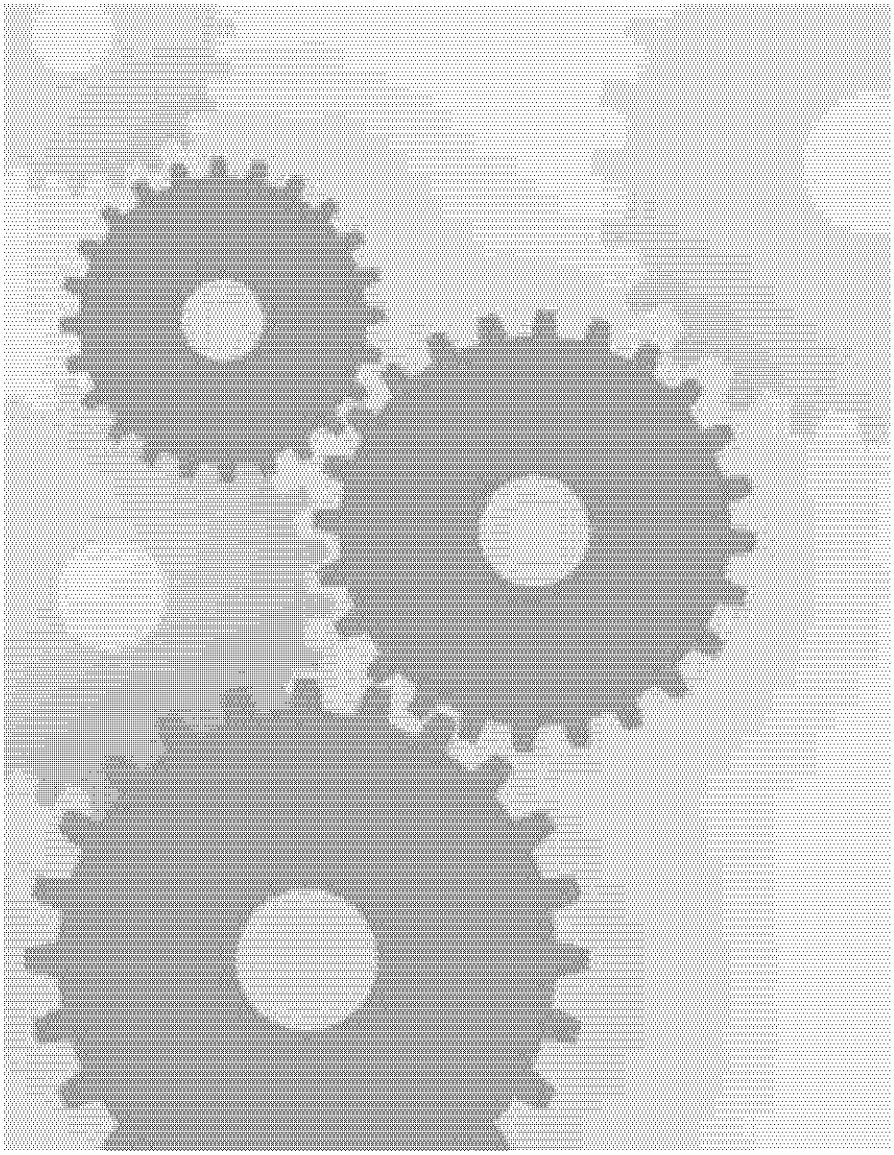
The Decision Support System Literacy Task Force hopes you find the *Forum Guide to Decision Support Systems: A Resource for Educators* useful, and that it helps improve data-driven decisionmaking in schools, school districts, and state education agencies across the nation.

Tom Ogle
Task Force Co-Chair
Missouri Department of
Elementary and Secondary Education

Raymond Yeagley
Task Force Co-Chair
Northwest Evaluation Association

Forum Review Procedures

Task force members review all products iteratively throughout the development process. Documents prepared, reviewed, and approved by task force members undergo a formal public review. This public review consists of focus groups with representatives of the product's intended audience, review sessions at relevant regional or national conferences, or technical reviews by acknowledged experts in the field. In addition, all draft documents are posted on the Forum website prior to publication so that any interested individuals and organizations can provide feedback. After the task force oversees the integration of public review comments and reviews the document a final time, all publications are subject to examination by members of the Forum standing committee sponsoring the project. Finally, the entire Forum (approximately 120 members) must review and vote formally to approve the document prior to publication.



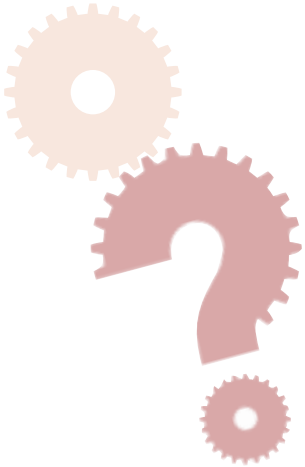
Introduction

Many education stakeholders want access to more data to help them decide how best to operate, manage, and evaluate our schools. But they do not want just any data–they want better data. They want real-time data they can use to run their schools more efficiently today; up-to-date information that permits them to compare school inputs, processes, and outcomes during the current grading period; and longitudinal information that enables them to anticipate their schools' needs in the future. In other words, they want data to be useful, accurate, well organized, and readily accessible to those who need it to make decisions about the operation and management of the education enterprise.

"Data-driven decisionmaking" is critical to many organizations across the nation, including schools, school districts, and state education agencies. In an education setting, it means that pedagogical and operational choices are to be informed by accurate, relevant information available in time to influence decisionmaking. To do this, however, raw data from disparate sources must be accessed, integrated, compiled, and distilled into useful information in a timely manner. This task may best be accomplished by a specific class of computer information systems called "decision support systems (DSS)." Many education organizations trying to get the right data into the hands of the right decisionmakers at the right time have concluded that investing in such a system might be the best solution for their information management needs.

Investing in a decision support system promises numerous benefits, which in many circumstances may outweigh the costs, but it is nonetheless a major decision. Purchasing a decision support system represents a significant financial and operational commitment. Some of the costs are related to hardware and software, but there are other expenses as well–potentially including redesigning the organization's data architecture, changing data collection procedures, and upgrading system security. Initial and ongoing stakeholder training and support may also be necessary.

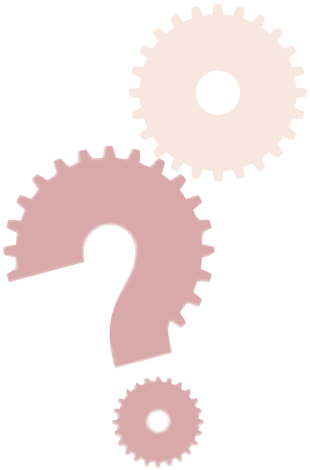
Decision support systems are inherently complex in terms of both their data management and technology architecture. This natural complexity is amplified by sometimes conflicting information coming from the technology industry. Some of this confusion stems from genuine disagreement over definitions and an inability to clearly delineate related concepts, such as data marts and data warehouses. In other cases, vendors selling one technology solution may not be able, or willing, to accurately describe the similarities and differences between their product and other available systems. Therefore, many decisionmakers in education organizations (who are educators rather than technology specialists) may find it difficult to obtain the reliable and objective information they need to better understand decision support systems and determine how they might be used most effectively in education organizations.



The *Forum Guide to Decision Support Systems: A Resource for Educators* was developed to remedy this lack of reliable and objective information about using decision support systems in education organizations. In other words, this guide was written to "educate the educators" about decision support systems. More specifically, this document addresses the following broad questions:

Part I. What is a Decision Support System?

- What is a decision support system?



- How does a decision support system differ from a data warehouse and a data mart?
- What types of questions might a decision support system address when used in an education organization?

Part II. Components of a Decision Support System

- What components, features, and capabilities commonly comprise a decision support system?
- How does each broad category of these components and features contribute to system operation?

Part III. Developing a Decision Support System

- How does an education organization buy or develop a decision support system?
- How are stakeholders trained to use a decision support system?

A decision support system might be used in a great many different ways in the numerous schools, districts, and state education agencies across the nation. This wide range of applications and settings creates substantial variation in the features and capabilities of decision support systems; however, a common core of concepts and models are associated with most decision support systems. This document emphasizes those commonalities to help educators better understand these fundamental concepts and capabilities prior to making significant financial investment in a decision support system.

Part 1
WHAT IS A DECISION SUPPORT SYSTEM?

Questions to be addressed:

- What is a decision support system?
- How does a decision support system differ from a data warehouse and a data mart?
- What types of questions might a decision support system address when used in an education organization?

Although many outside the systems technology and data management communities may not be very familiar with the term "decision support system," the concept was initially conceived as early as the mid-1960s.[1] Forty years later, these valuable systems are used in private and public sector organizations around the world, although they are still not well understood by those without highly technical training. While describing all types and permutations of decision support systems is beyond the scope of this publication, this document can help readers better understand what decision support systems are, how they are configured, how they operate, and how they can be implemented in an education organization.

Defining the Concept

A "decision support system" may be defined in many ways. Some definitions emphasize hardware and software components; others focus primarily on function (i.e., fulfilling the information needs of decisionmakers); while a few even describe user interfaces, job functions, and data flow. As such, competing yet complementary definitions of decision support systems include:

- Decision support system: An interactive software-based computerized information system intended to help decisionmakers compile useful information from raw data, documents, personal knowledge, and business models to identify and solve problems and to make decisions.[2]
- Decision support system: An interactive computerized system that gathers and presents data from a wide range of sources to help people make decisions. Applications are not single information resources, such as a database or a graphics program, but rather the combination of integrated resources working together.[3]
- Decision support system: A cohesive and integrated set of programs that share data and information and provide the ability to query computers on an ad-hoc basis, analyze information, and predict the impact of possible decisions.[4]

A decision support system is clearly not an application that simply manipulates data or supports decisionmaking. For example, an enhanced user interface that permits querying and analysis of a single database is not a decision support system; nor is a spreadsheet application with basic analysis and advanced "if/then" planning features. Even a database management system (DBMS) that permits a user to select and analyze data within a single database for reporting and analysis would not qualify, because it does not integrate multiple databases. Rather, a decision support system is intentionally and explicitly more comprehensive, and is designed specifically to enable users to support problem solving and decisionmaking by compiling information from disparate sources of raw data.

A robust definition of a decision support system should encompass: (1) users who understand what the data mean and how they can be accessed with a (2) technology system (hardware, software, and user interfaces) that manipulates (3) a data system (integrating data from multiple sources) explicitly for the purposes of (4) a decisionmaking system (user-driven within an organization). While not a formal definition, this description was developed for this publication to stress multiple emphases on user skills, technology tools, data quality, information use, and organizational management encompassed by true decision support systems. Such a description incorporates technology tools for managing, analyzing, communicating, and using data; an understanding of data within the system and the implications of the use of those data; and an intention by decisionmakers to employ information for the purpose of planning and action within an organization.

For the purpose of this document, a "decision support system" is defined as a cohesive, integrated hardware and software system designed specifically to manipulate data and enable users to distill and compile useful information from disparate sources of raw data to support problem solving and decisionmaking.

Data Warehouse and Data Mart

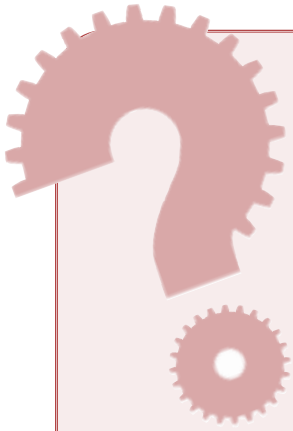
Two terms are often confused with "decision support system": "data warehouse" and "data mart." While superficially similar in both terminology and concept, they are in fact different and should not be used interchangeably.

A "data warehouse" is a central repository for all, or a significant portion, of the data an enterprise collects. Data "warehousing" emphasizes storing data from diverse sources, but it does not generally concern itself with the end user as a decision support system would.[5] To be more specific, data warehousing serves the functions of "querying and reporting" large sets of data, rather than "querying and analyzing" data. Moreover, the primary purpose of a data warehouse is to provide access to historical or transactional data in their basic format (e.g., tables), not to distill data into a format that encourages or even permits in-depth analysis.[6]

A "data mart" is very similar to a data warehouse in that it is a repository for data; however, data marts are limited in scope to a subset of an organization's information, as delineated by subject, function, utility, or user group. Like a data warehouse, a data mart does not offer the sophisticated data analysis and reporting capabilities of a decision support system.

Transforming Education Decisionmaking

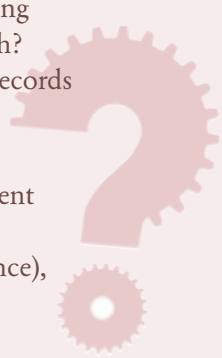
Utilizing a decision support system is a proactive way to use data to manage, operate, and evaluate education institutions. Depending on the availability and quality of the underlying data, such a system could address a wide range of questions by distilling data from any combination of education records systems. The following examples illustrate a few of the many questions a robust decision support system could potentially address.



Questions about the classroom.

- Do students who have teachers with degrees in mathematics perform better on math assessments than students whose teachers have degrees in other areas?
  • To answer this question, the system might compile data from staff records (degree type), school records (classroom teaching assignment), and student gradebook systems (assessment results).
- Are all students in the fourth grade progressing at the same rate, or are the students who had a specific third-grade teacher doing better than the others?
  • Answering this question requires data from school records (classroom teaching assignment) and student gradebook systems (academic progress).
- Are the students who receive Title I services progressing at the same rate as those who are not?
  • This question requires data from program records (Title 1 participation) and student gradebook systems (academic progress).
- Are Hispanic students progressing at the same rate as students from other ethnic backgrounds?
  • The answer to this question requires data from student information systems (race/ethnicity) and student gradebook systems (academic progress).

Operational questions that extend beyond the classroom.



- Are students in Supplemental Educational Services (SES) programs showing academic growth on large-scale assessments, improved attendance, or both?
  • To answer this question, the system might compile data from program records (SES participation), student gradebook systems (academic growth), and student information systems (attendance).
- Does a reduction in staff injuries correlate with a district's staff development activities or other safety measures?
  • Answering this question requires data from staff records (injury occurrence), human resources systems (staff training topics and participation), and school facilities systems (measures to improve safety).
- Are there fewer veteran teachers at lower-performing than at higher-performing schools?
  • To answer this question, the system might compile data from staff records (years experience), school records (school teaching assignment), and student gradebook systems (assessment results).

Policy questions at the state level.

- Do one district's students perform at a higher level than those in other districts with similar demographics and per pupil expenditures?

• To answer this question, the system might compile data from district-level records from across a state, including student demographic records (race/ethnicity, socio-economic status, etc.), school finance systems (per pupil expenditures), and student gradebook systems (academic performance).

- Is there a correlation between the amount of school district funds dedicated to early childhood education and student performance?
  • This question requires data from state-level records, including school records (student performance), program systems (early childhood education participation), and finance systems (fund allocation).
- Do urban and rural districts have fewer highly qualified teachers than suburban districts?
  • Answering this question requires data from staff records (highly qualified status) and geographic information systems (urban/suburban/rural designations).
- Are more minority students passing higher-level algebra courses in middle schools?
  • To answer this question, the system might compile state-level data from student information systems (race/ethnicity), student gradebook systems (class performance), and school records (school type).

Summary

Decision support systems are becoming increasingly important information management tools in education organizations. They are already being used effectively by many schools, districts, and state education agencies across the nation. Depending on their configuration, these systems can be powerful tools for addressing a wide range of questions about student performance, classroom management, organization-wide operations, and state-level policymaking.

Notes

[1] Power, D.J. *A Brief History of Decision Support Systems.* Version 2.8 retrieved April 18, 2006 from *http://dssresources.com/history/dsshistory.html.*

[2] Adapted from *Decision Support Systems–DSS (definition),* Information Builders. Retrieved April 20, 2006 from *http://www.informationbuilders.com/decision-support-systems-dss.html.*

[3] Adapted from *Webopedia.* Retrieved April 20, 2006 from *http://www.webopedia.com/term/d/decision_support_system.html.*

[4] Adapted from *PCMag Encyclopedia.* Retrieved April 20, 2006 from *http://www.pcmag.com/encyclopedia_term/0,2542,t=DSS&i=42054,00.asp.*

[5] Adapted from *WhatIs.com.* Retrieved April 20, 2006 from *http://whatis.techtarget.com/definition/0,289893,sid9_gci211904,00.html.*

[6] Adapted from Greenfield, L., *A Definition of Data Warehousing (2006).* The Data Warehouse Information Center. Retrieved April 19, 2006 from *http://www.dwinfocenter.org/defined.html.*

Part II
COMPONENTS OF A DECISION SUPPORT SYSTEM?

Questions to be addressed:

- What components, features, and capabilities commonly comprise a decision support system?
- How does each broad category of these common components and features contribute to system operation?

Common Components of Decision Support Systems

Describing a single model that would apply to every decision support system used in all education settings is challenging if not impossible. Still, most of the systems used in an education environment have some features in common. The following description is based on these commonalities.

In general terms, decision support system components commonly include:

Data Quality: The Foundation of Any Decision Support System

1. Data collection

Hardware, Software, and Data Management Processes

2. Hardware, networks, and operating systems
3. Underlying data sources
4. Extract, transform, and load (ETL) process
5. Data warehouse or data aggregator
6. Analysis and reporting tools
7. User dashboard *(see figure 1)*

A DSS does not make decisions; it only supplies information, presented efficiently, to help staff make decisions.

Data Quality: The Foundation of Any Decision Support System

1. Data collection

Data are the foundation of any decision support system. High quality data are useful (relevant to decisionmaking), valid (accurately measured), reliable (reproducible), and timely (available in time to influence decisionmaking). Data systems that produce quality data often emphasize:[1]

- appropriate data collection schedules;
- rigorous verification and documentation requirements;
- thorough validation procedures;
- clear, accessible, and customized coding instructions;

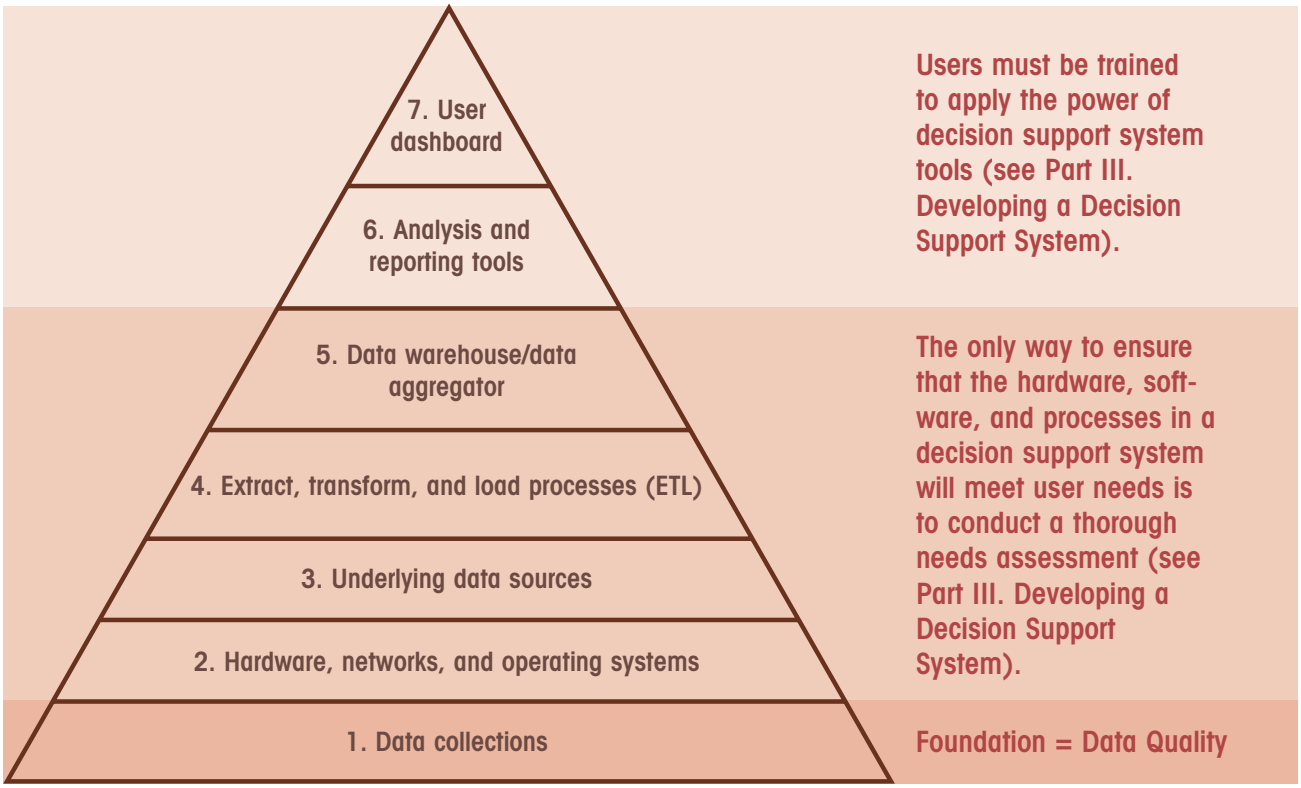


Figure 1. Although no single model applies to all decision support systems, both the elements and their relationship to one another, as shown in figure 1, reflect features common to many systems that would be used in education organizations. For information about conducting a needs assessment, training users, and other critical elements to planning a decision support system, see Part III. Developing a Decision Support System.

- effective training and support programs;
- accurate and consistent data entry;
- automated data transfer processes; and
- consistently applied terminology (a data dictionary) and rules for data collection and maintenance (metadata).

Perhaps the single most important mechanism for collecting and maintaining high quality education data is the consistent application of standard terminology and business rules throughout the organization. Most institutions with sound data systems have a single, exhaustive data dictionary that is the definitive source for data term usage expectations; data definitions; and other attributes typically associated with data elements, including field lengths, code lists and definitions, formats for each given type of data (e.g., mm/dd/yyyy format for dates), and any restrictions on values or value ranges (e.g., "age" must be a value between 1 and 99).[2] A corollary to the use of a data dictionary is the explicit link between each piece of data, the data dictionary, and other information that describes the context of the collection or use of the item. This "metadata," which is most simply defined as "data about data," might include, for example:

As with any data-based system, the quality of information available from a decision support system depends on the quality of the data originally entered into the system.

- when data were collected;
- how data were collected (e.g., survey name, definitions, and instructions);
- how data were originally formatted (and any subsequent changes);
- when a record was accessed or modified;
- data dictionary references (e.g., code lists, field length restrictions, and technical parameters);

- who "owns" the data item (i.e., which individual or office has the final authority to change a definition or attribute and coordinate appropriate implementation across the rest of the organization); and
- other information needed to appropriately use or interpret the data.

Hardware, Software, and Data Management Processes

2. Hardware, networks, and operating systems

A decision support system is not a single piece of technology, such as a database, file server, or network. Rather, it is a system for incorporating and integrating disparate data sources to better allow decisionmakers to access and compile data in a useful format. The technology associated with a decision support system depends on the organization's preexisting technologies and data architecture (i.e., how data are currently stored and accessed), as well as the organization's technical and functional requirements identified during planning efforts (see part III). In general, most decision support systems will include the hardware, networking technologies, and operating systems necessary for supplying and supporting databases and/or servers; a user interface with mechanisms for accessing, manipulating, and transferring data; and some type of repository for temporarily or permanently storing data. Key technical requirements often revolve around issues such as accessibility, processing and transfer speed, scalability, interoperability, cost-effectiveness, and security.

3. Underlying data sources

The data used to make decisions in education organizations come from many disparate sources, including school business office databases, attendance sheets, student assessment results, athletic team rosters, human resources files, curriculum frameworks, and school bus routing systems. External sources such as Census data and county zoning data (e.g., geographic information systems) might also be needed, for example to select a building site for a new school. Moreover, some data may be entered directly into school computer systems after collection by traditional paper-and-pencil methods, while other data are downloaded from databases at district administrative offices, state education agencies, and national testing companies. Another complication is the fact that some data may reflect up-to-the moment, real-time events, while others represent archived records from previous decades.

The distributed nature of these data sources presents three challenges that must be addressed before a decision support system can effectively produce useful information:

- Accessibility
- Ownership
- Interoperability

Accessibility. Decision support systems are designed specifically to gather and present data from a wide range of sources. If the system does not provide access to all available data, or at least all data to which the user has access privileges, it is not a decision support system, but rather a data mart or data warehouse. As discussed above, these are powerful and useful information management tools but not fully integrated decision support systems.

Ownership. Within an organization, different groups or departments may claim to "own" databases. For example, a program office may collect data on a particular topic each year, or a school may manage its own operational records (daily attendance, class schedules, teaching assignments, etc.). While data quality is generally higher when individuals feel personally responsible for "their" data, certain decisions about how data are collected and maintained should be coordinated across the entire organization. Consider, for

example, who has the authority to change the format, definition, or content of the "Student Address" data element; both the transportation director who uses student addresses to plan busing service and the school secretary who uses the same addresses to send mail may have legitimate "claims" to this element. Unless ownership authority is well defined, the organization risks having multiple users modifying formats, definitions, and code lists independently. This may create redundant versions of the same information, and each version may be defined, coded, or formatted slightly differently. Such parallel systems lead to inefficient use of hardware (storing multiple data sets), wasted staff time (having to code one change of address separately in multiple places), unnecessarily restricted or delayed access (extracting from multiple locations), and decreased processing speeds (having to translate multiple sets of the same information).

Interoperability. The term "interoperability" refers to the ability of a system or component to work with other systems or components. In a data system, interoperability denotes the ability of two or more systems to exchange and integrate data in the absence of additional steps to translate the data from one system to the other. Ideally, even separate databases within a single organization will be designed so that data can be exchanged, manipulated (e.g., aggregated), and reported seamlessly and without conflict. Unfortunately, different information systems within the same organization may not always be interoperable–one database may have slightly different technical specifications or data formats than another.

For example, an item as straightforward as "Gender" may be difficult to reconcile when it is maintained in slightly different formats in different databases; e.g., "Male or Female," "M or F," and "1 = Male and 2 = Female" are all valid but different entries for "Gender." Table 1 illustrates different formats for the same data element, "Name." Although the data may be entered correctly in each database, these different formats present a problem when, for instance, a query tool searches all underlying databases in the organization. Given that a decision support system is intended to integrate all data systems in an organization, a process of "extract, transform, and load" (ETL) is often required when this occurs (see below). Many organizations choose to introduce ETL processes because establishing interoperability between old (legacy) and new systems is often difficult, expensive, and time-consuming.

DATABASE	NAME DATA	FORMAT	COMMENT
Student information system	Mary Ruth Smith	Full name	One field
Financial information system	Smith, Mary (R)	Last name, first name (middle initial)	Two fields
Human resources system	Mary, Ruth, Smith	First name, middle name, last name	Three fields

Table 1. "Name" data formatted correctly, but differently, in three separate databases.

4. Extract, transform, and load (ETL) process

The extract, transform, and load (ETL) process is necessary when source data in a decision support system reside in separate, non-interoperable databases *(see table 1)*. As the name implies, ETL is a three-stage process designed to move data from legacy source systems into an interoperable format in the decision support system. In the first step, an "extract" function reads from a specified database and pulls out the desired data. In step two, a "transform" function uses predetermined business rules to convert the extracted data into a format that is interoperable (see above) with other system data. Finally, in step three, a "load" function moves the edited and cleaned data to a database repository

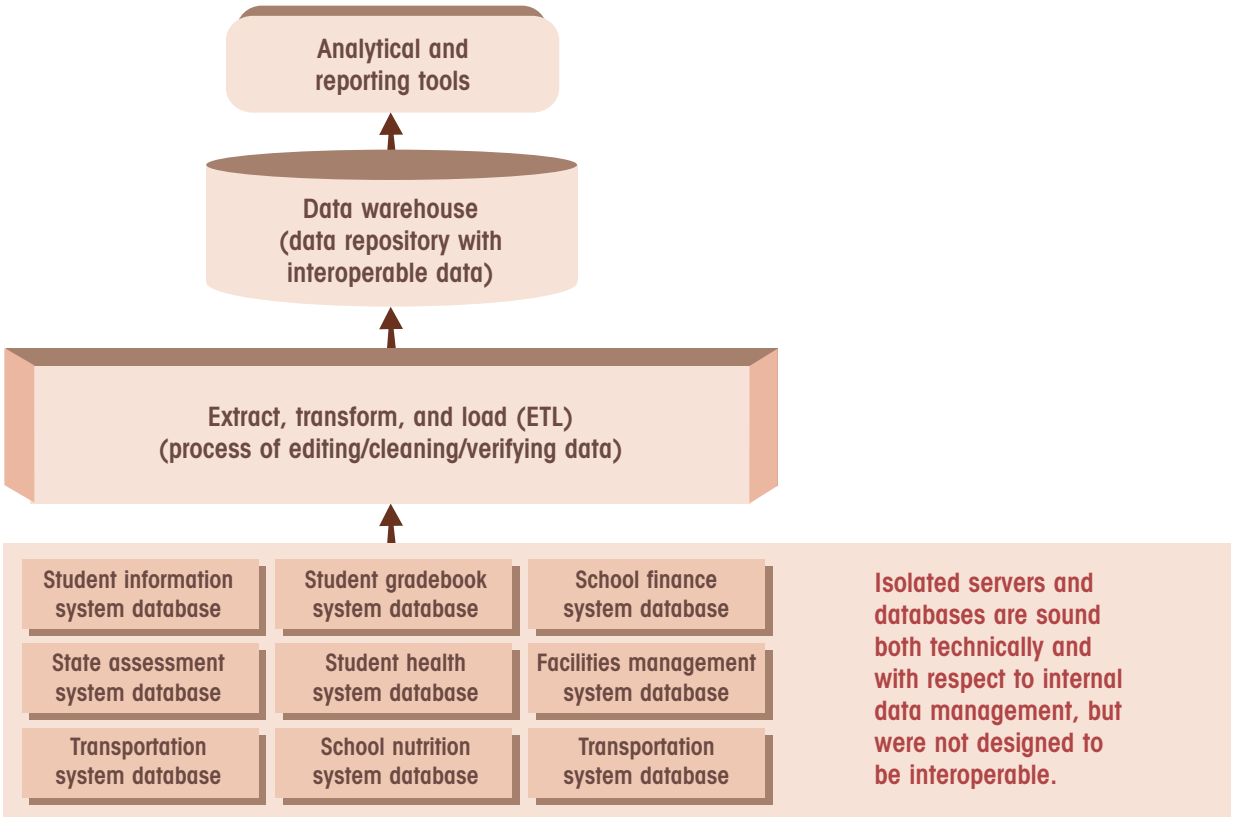


Figure 2. Separate databases can be well designed, both technically and with respect to data management practices, but unless they are designed to be interoperable, extracted data will need to be transformed (edited, cleaned, and verified) before they can be integrated and accessible in a decision support system data repository.

(often called a data warehouse) within a decision support system *(see figure 2)*. In some cases, the ETL process includes a fourth step, referred to as a staging area (or, informally, a "sandbox"). In this context, a "sandbox" is a testing tool that allows system administrators and users to test the loaded data in a practice setting before final transmission to the "live" environment. By allowing users to double check data sets in a preproduction setting, where changes can be made more easily, costly and time-consuming data-cleaning activities in the actual system are often minimized.

5. Data warehouse or data aggregator

As seen above, user queries may require compiling data from disparate sources within (and sometimes outside) an organization. When these sources are not interoperable, the ETL process must be used to edit, clean, and verify data that is then stored in a data warehouse for analytical and reporting use *(see figure 2)*. Conversely, if different databases are initially designed (or later adapted) to be interoperable, the need for an ETL process is greatly reduced. When this is the case, the ETL step can be replaced with a "data aggregator," a tool that simply locates and compiles queried data rather than editing, cleaning, and verifying it *(see figure 3)*.

6. Analysis and reporting tools

Analysis tools. An "analysis tool" is basically an instrument that applies business rules or other logic to data in order to derive meaning. This includes time series analysis, cost allocations, data mining, and other user-driven manipulation and investigation. Analysis

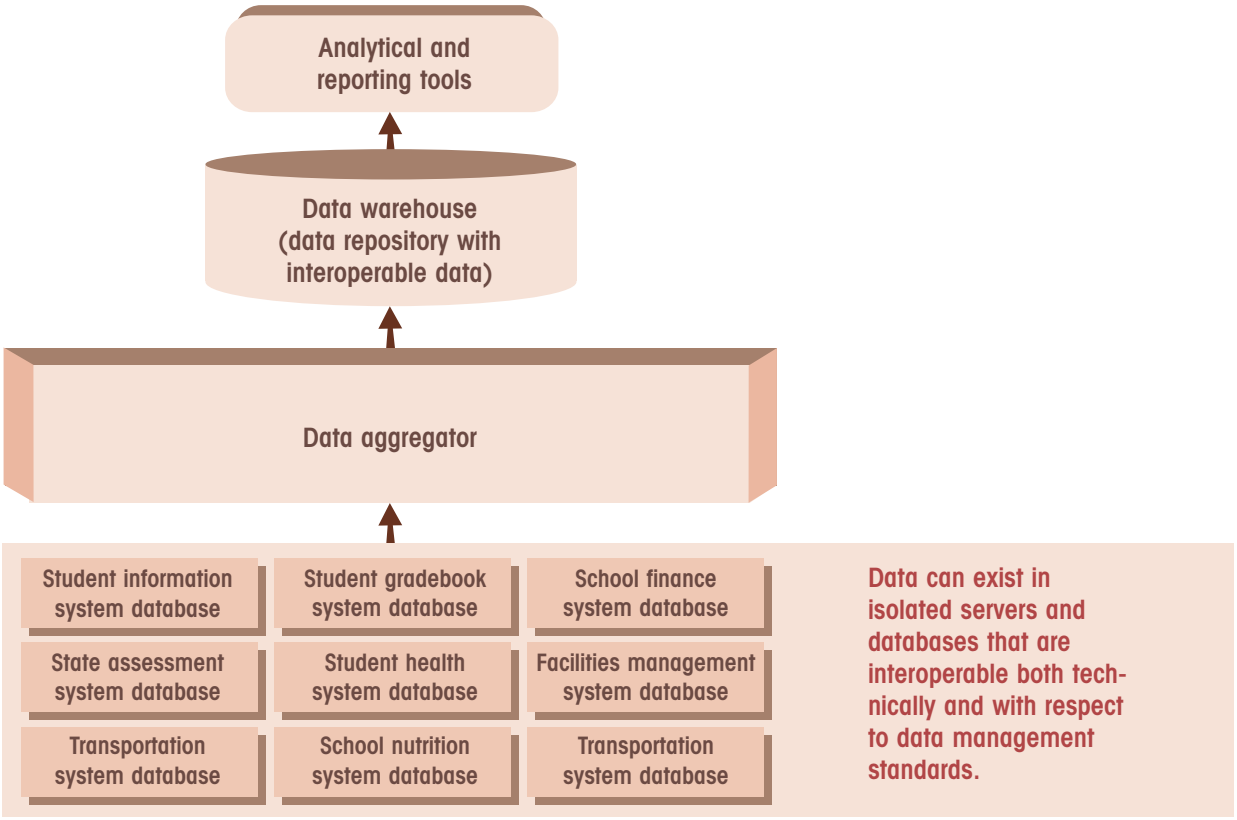


Figure 3. The need for the ETL process can be greatly reduced if separate data sources are designed to be interoperable—that is, actively managed with a single data dictionary, metadata convention, and data ownership standard. When this occurs, a data aggregator may be the correct tool for moving data from multiple, interoperable data sources into a decision support system's data warehouse.

tools are available in many software applications, including spreadsheets, databases, and other stand-alone programs. In a decision support system environment, however, analysis tools are particularly powerful because they rely on On-Line Analytical Processing (OLAP) technologies. OLAP tools are applications that permit users to browse, query, analyze, and summarize large amounts of data in an efficient, interactive, and dynamic way. Such a tool is a useful component of decision support systems because of the multi-dimensional nature of large data sets. For example, consider a spreadsheet with rows for classrooms and columns for average test scores by subject area; with OLAP, those same two-dimensional data cells (the intersections of rows and columns) can be organized into layers by year, adding a third dimension of time *(see figure 4)*. When OLAP tools generate this type of three dimensional data, the output is sometimes referred to as an "OLAP cube." OLAP cubes permit the manipulation of data between dimensions by relatively simple, point-and-click user interface rather than complex statistical programming; for example, the data can be "pivoted," or presented using any of the three dimensions as the primary unit of analysis. The ability to manipulate data in multiple dimensions improves data analysis and reporting capabilities, making OLAP cubes invaluable for data mining, data management, and trend analysis–and powerful analytical components of decision support systems.[3]

Reporting tools. Robust reporting tools are a major element of any decision support system. Presenting information in multiple formats (as a blend of text, tables, and graphics)

and in multiple dimensions (changing an axis to present information more clearly, as discussed above) sometimes clarifies the meaning of the data. Unlike a data warehouse or database, which both focus on data storage, a decision support system often includes reporting tools that permit a user to easily:[4]

- place headings, titles, and explanatory information within charts, tables, and other derived figures;
- add borders and shading to clarify and highlight important information and groupings;
- modify font size and style to emphasize points;
- move, edit, or delete data, text, and graphics in final reports;
- produce a wide range of figures, including bar graphs, pie charts, bar and line graph combinations, multiple axis graphics, and scatter plots;
- export data in various formats (e.g., ASCII, Excel™);
- generate reports in various formats (e.g., html, PDF™, e-mail, paper); and
- include legends, citations, explanations, and other information.

A decision support system's reporting functions must serve a wide range of users–including novices and users with expert analytical capabilities. To accommodate this, most systems offer two primary classes of reporting tools: (1) predefined (static) reports that require little system expertise and are ideal for users with typical information needs; and (2) dynamic (ad-hoc) report-generating capabilities that require greater understanding of both the data and the querying technology, but allow users to investigate more complex questions.

Predefined reports. Some types of data requests are quite common: How many students are enrolled this year? How many students graduated last year? What percentage of students took advanced math courses in the past five years? Because these and many other data requests are quite common in education settings, they can be anticipated and are often preprogrammed, in predefined reports. These types of reports are especially effective tools for users who require basic and predictable information. The more predefined reports a decision support system offers, the more needs it can serve without requiring stakeholders to become specialists at querying the system.

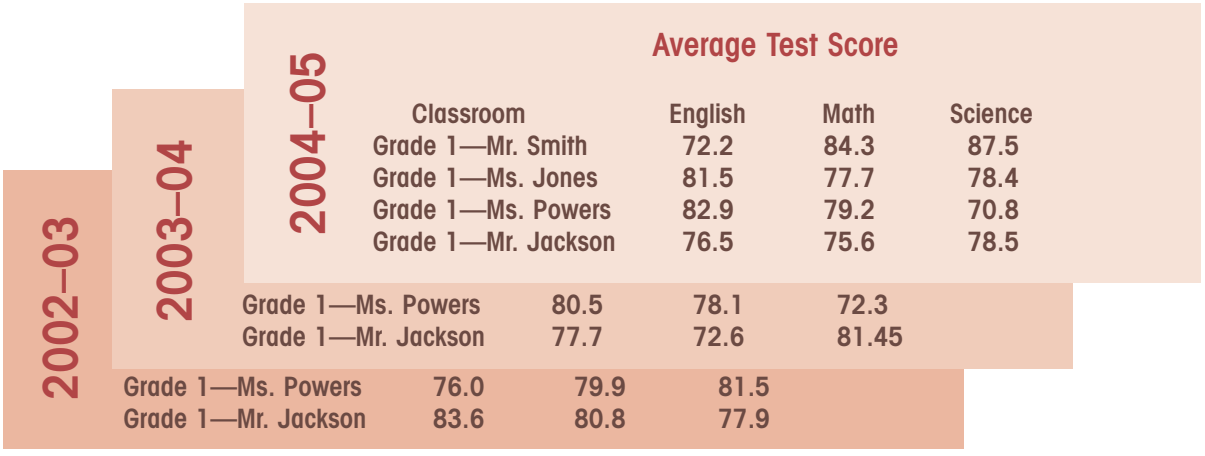


Figure 4. Illustration of multidimensional data. The three dimensions shown in this example include rows (classrooms by grade level and teacher name), columns (average test scores by subject area), and layers (years over time). OLAP cubes are multidimensional in nature and are designed to efficiently display and manipulate multidimensional data.

Ad-hoc reports. Whenever existing, predefined reports cannot provide an appropriate response to a query, users may be able to customize their request and generate an "ad-hoc" report. In the context of querying, "ad-hoc" refers to a data request that is tailored to meet the specific needs of an individual user. By definition, ad-hoc queries are not readily predictable and cannot, therefore, be preprogrammed by system managers. Clearly, users who require ad-hoc reporting tools will probably need a more sophisticated understanding of how to use querying tools.

7. User dashboard

In a decision support system, a "dashboard" serves as a user interface that both presents information and enables a user to access or compile new data by means of a series of "gauges and dials." Unlike a car's dashboard, which is simply a portal to view data about the vehicle's condition and operation, a user dashboard in this context is an actively managed and integral component in the continuum of data collection, validation, analysis, reporting, and decisionmaking that constitute a decision support system.

Dashboards are most effective when they are customized to reflect the specific needs of each user group. This requires the collaboration of: (1) stakeholders to communicate

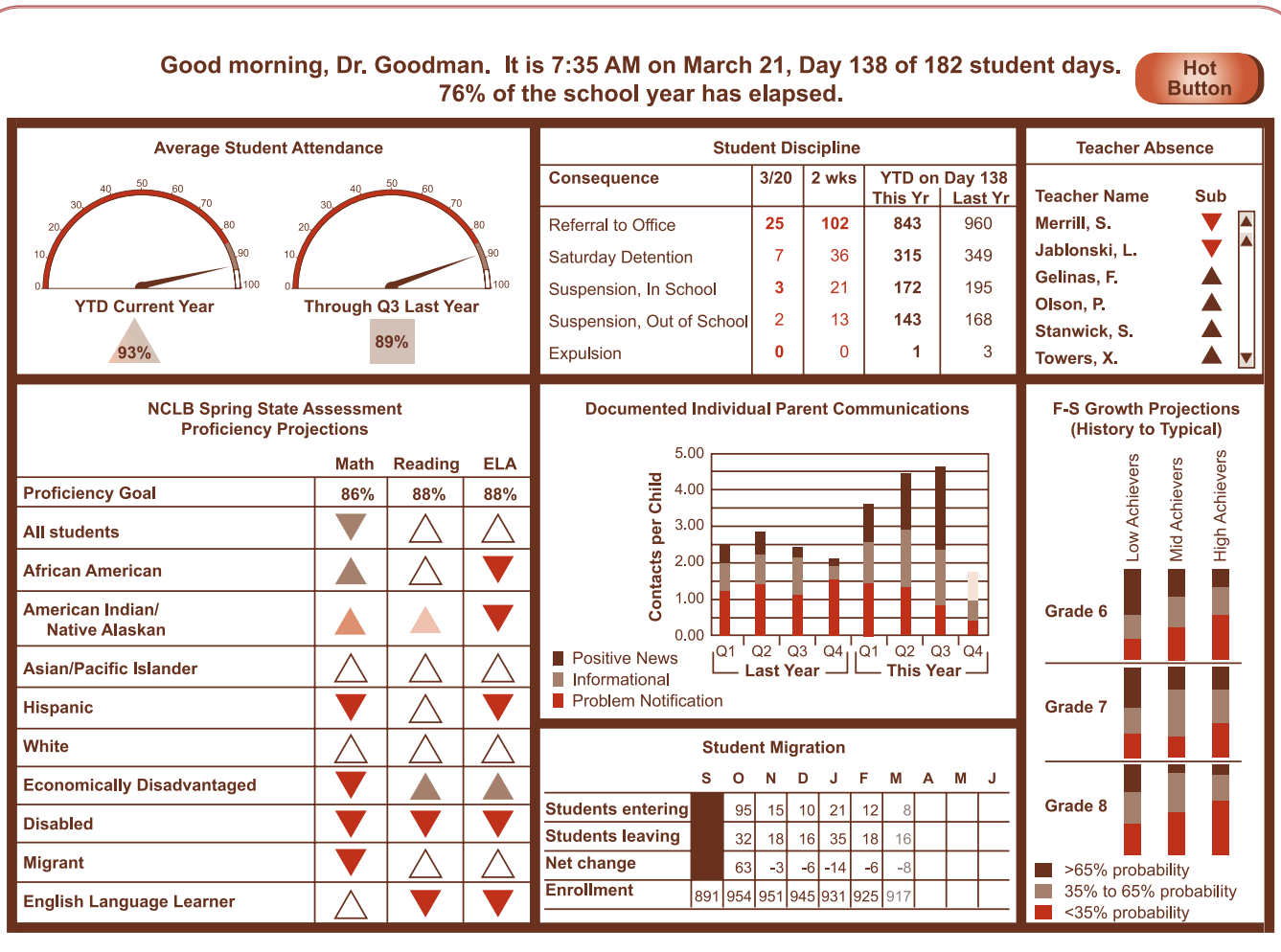


Figure 5. Just as a car's "dashboard" indicates key metrics for operating the vehicle (fuel balance, speed, emergency warnings), a decision support system dashboard indicates the current status of an education organization, such as daily attendance, recent disciplinary incidents, teacher absenteeism and substitute needs, recent assessment scores, parent feedback, and student population changes. Decision support system dashboards generally are customized by user type to ensure they display the information needed to inform decisionmaking.

their information needs (i.e., what questions they regularly ask to meet their responsibilities); (2) data specialists to determine what data are needed to answer these questions; and (3) technical staff to translate input from stakeholders and data specialists into a dashboard interface. For example, a school principal's dashboard might display data about daily attendance, recent disciplinary incidents, teacher absenteeism, substitute teacher needs, recent assessment scores, parent feedback, and student population changes *(see figure 5)*.

Summary

Although describing a single model that would apply to all decision support systems is challenging, some features are common to most of the systems used in education organizations. Common features include data collection activities; hardware, networks, and operating systems; underlying data sources; extract, transform, and load (ETL) processes; a data warehouse or data aggregator; analysis and reporting tools; and a user dashboard.

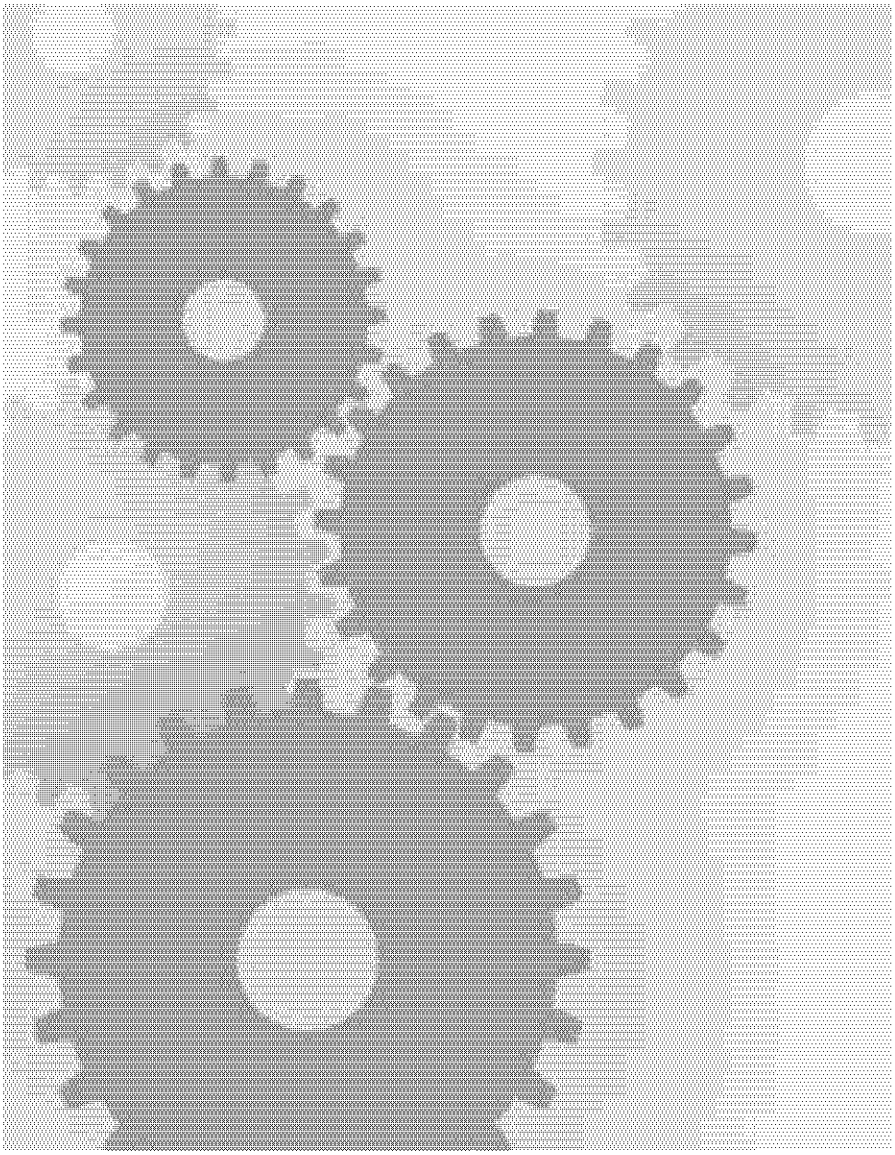
Notes

[1] Adapted from *Accounting for Every Student: A Taxonomy for Standard Student Exit Codes* (NFES 2006–804). National Forum on Education Statistics. (2006). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Available at *http://nces.ed.gov/forum/pub_2006804.asp*.

[2] Adapted from *Forum Guide to Building a Culture of Quality Data: A School and District Resource* (NFES 2005–801); National Forum on Education Statistics; (2004); U.S. Department of Education; Washington, DC: National Center for Education Statistics. Available at *http://nces.ed.gov/forum/pub_2005801.asp*. Material was also adapted from the *NCES Handbooks Online*, retrieved April 27, 2006 from *http://nces.ed.gov/programs/handbook*.

[3] For more information about OLAP cubes, tools, and logic, see Dramowicz, K., Creating and Manipulating Multidimensional Tables with Locational Data Using OLAP Cubes, *Directions Magazine* (January 15, 2005). Retrieved April 27, 2006 from *http://www.directionsmag.com/article.php?article_id=733&trv=1*.

[4] For more information about standards of report display and presentation, see appendix C of the *Forum Guide to Education Indicators* (NFES 2005–802); National Forum on Education Statistics; (2005); U.S. Department of Education; Washington, DC: National Center for Education Statistics. Available online at *http://nces.ed.gov/forum/pub_2005802.asp*.



Part III
DEVELOPING A DECISION SUPPORT SYSTEM

Questions to be addressed:

- How does an education organization buy or develop a decision support system?
- How are stakeholders trained to use a decision support system?

Developing a decision support system is not a technology project–it is a data project driven by business needs and supported by technology tools. Many methods for planning this type of initiative have been published; most contain the same general steps, although the tasks may be described differently. In general, these steps are as follows:

1. Define the task, conduct a needs assessment, establish technology and functional requirements, and describe current resources.
2. Evaluate defined needs relative to current capabilities.
3. Perform cost-benefit analysis and select a solution that best meets the goals of the initiative.
4. Purchase (or develop) and install the selected solution.
5. Secure technology and information based on findings from the risk assessment.
6. Plan for ongoing system maintenance and support.
7. Train users to maximize the utility and efficiency of the new system.
8. Integrate the resources and processes into daily routines and long-range planning.

Once planners have determined the organization's needs and requirements, they can prioritize those needs based on cost-benefit analysis, and document the results for inclusion in a Request for Proposals (RFP) (see appendix).

While addressing each of these steps in detail is not in this document's purview, some points are particularly relevant to education organizations undertaking the development of a decision support system.[1]

Conducting a Needs Assessment

A "needs assessment" is an evaluation of all the tasks and functions an organization should be capable of performing. Most data and decisionmaking needs arise from "users," the people who "use" the organization's data to decide how to manage, operate, and evaluate a school, district, or state education agency. One of the challenges faced by planners is to create a decision support system that meets the needs of every different type of user, collectively referred to as "stakeholders." In an education setting, stakeholders

might include teachers, principals, school secretaries, program managers, superintendents, board members, local policymakers (e.g., the mayor or county administrator), students, parents, and other community members. The list of potential stakeholders is long and varies by organization; to be thorough, anyone who uses data for decisionmaking should be considered a candidate for participating in the needs assessment, regardless of data or technical expertise.

While soliciting input from every stakeholder is impractical, representatives from each stakeholder group should be included in the planning stage. When decisionmakers are not thorough in this part of the planning, the final product often has missing features, redundant components, inadequate support after deployment, or stakeholders whose information needs are not met by the new system. In short, the need for good planning cannot be overstressed: correcting problems can be much more expensive and time-consuming than planning for contingencies in the first place.

Data Security Planning

In education organizations, the underlying data in a decision support system often include private information about students and staff members (e.g., contact information, health records, assessment results). Some of these data are protected by state and federal privacy laws, such as the Family Educational Rights and Privacy Act (FERPA) and the Health Insurance Portability and Accountability Act (HIPAA). Education organizations should enact strong policies and procedures to ensure compliance with state and federal regulations governing the privacy and confidentiality of personal information.[2] Relevant privacy laws include:

Information security policies are necessary in all education organizations, but a decision support system's power to increase data access amplifies their importance.

- The Family Educational Rights and Privacy Act (FERPA)
  The Family Educational Rights and Privacy Act (FERPA) protects all personally identifiable information in student records, except for items designated as "directory information." Under FERPA, even directory information (e.g., student name) may be restricted if parents or legal guardians object to its release for their own children.[3]
- The Health Insurance Portability and Accountability Act (HIPAA)
  The Health Insurance Portability and Accountability Act (HIPAA) protects the privacy of medical information for students, employees, and others in the data system. While HIPAA protects fewer kinds of records, it is much more restrictive than FERPA with respect to how data may be collected, stored, and shared.[4]
- The Individuals with Disabilities Education Improvement Act (IDEA)
  The Individuals with Disabilities Education Improvement Act (IDEA) adds additional layers of protection for certain information about students with disabilities.[5]

Although education data are usually reported as aggregates (i.e., data are compiled for groups rather than reported for individual students), values are sometimes so small or large that an enterprising viewer might be able to deduce personally identifiable information. For example, if a class has only one boy, reporting the "average" score of males on an assessment is the equivalent of sharing the boy's test results. Similarly, reporting that 99 percent of students in a school are eligible for free- or reduced-price meals is nearly the same as disclosing this data for each child in the school. Statisticians have developed sophisticated cell suppression techniques, which can be incorporated into reporting tools, to ensure that aggregate reporting does not violate privacy rights afforded to individual students.

The power to analyze data across multiple parameters using decision support system OLAP cubes (see part II) raises additional privacy issues because of their potential to generate smaller cell sizes than exist in other aggregated data. For example, disaggregat-

ing by race, poverty, and English-language status in a class of 30–which may be readily accomplished with a decision support system–may generate a report of three or four students per subgroup. Data stewards should work diligently to ensure they do not unintentionally violate confidentiality practices by providing data that, under common research scenarios, may generate personally identifiable information.

Fortunately, a properly configured decision support system can actually help protect data confidentiality. Role-based access rules are often used to protect confidential personal information; for example, parents may have access only to personally identifiable information about their own children, whereas teachers might have access to personally identifiable information about any student directly related to their official duties. With input from data and research specialists, information technology staff can create access rules in decision support systems that restrict the release of confidential information to non-authorized users. These protocols can be set to automatically suppress data that do not meet established criteria, thereby ensuring that users only view appropriately presented data.

User Training

Properly trained users are perhaps the most critical component of effective decision support systems. After all, people, not decision support systems, make decisions–the decision support system is only the tool that supports a decisionmaking process undertaken by users. In fact, the most difficult aspect of using a decision support system is not implementing the technology, but knowing what questions to ask, how to ask them, and how to interpret the answers (i.e., how to read the reports). Fortunately, users can be trained to understand the data and its limitations, as well as the system and its capabilities.

The best way to ensure that users know how to use the system and data appropriately is to train them. Initiating training activities prior to implementation prepares users for their introduction to a new system. Follow-up training once the system has gone "live" permits users to learn with actual data. Additional training sessions after users have had a chance to actually use the system are often effective in overcoming final obstacles to efficient use; this also gives users the chance to ask any questions that have come up while actually working with the system (and discuss any problems that occurred). While such training may seem expensive and time-consuming, the alternative is worse–expecting stakeholders to make good decisions, many of which also cost money and time, without the benefit of knowing how to use the system appropriately.

Differentiated professional development is a key to the successful implementation of a decision support system.

Many education organizations have found a "train-the-trainer" model to be effective as well as cost-efficient. In such an approach, staff who are intimately familiar with an organization's data, system, policies, and personalities are trained to use the system so that they can, in turn, customize and lead training sessions designed to meet the needs of other user groups. These training activities may be in person or online, depending on the school's location and technology resources.

Differentiated professional development. Because each stakeholder group may use a decision support system in a slightly (or substantially) different manner, developing separate training modules for each major type of user group often makes sense. For example, staff members who need to develop reports and graphic displays using OLAP cubes (see part II) will probably need very specialized training, as a thorough understanding of the system's capabilities is necessary to fully harness the power of the system. Alternatively, users who only need a few, predefined reports each grading period will likely work effectively with a less rigorous introduction to system capabilities–for example, a superintendent, business manager, or other administrator may not have the time to learn how to create ad-hoc reports, but they still need to understand any system dashboard created specifically for their use.

Ongoing professional development. After having had an opportunity to become familiar with the system, use it, and even make mistakes with real queries, users still require ongoing training. This not only allows for critical training points to be reemphasized, but also provides an opportunity to master complex issues over time. As in-house users learn the system, some may be able to serve as trainers in future training sessions.

Summary

The development of a decision support system should be driven by the business needs of the organization. Schools, districts, and state education agencies interested in purchasing or developing such a system may benefit from undertaking sound planning and implementation steps, as they would for any major information technology acquisition.

Notes

[1] These steps are described in greater detail in the *Forum Unified Education Technology Suite,* available online at *http://nces.ed.gov/pubs2005/tech_suite.*

[2] For more information about relevant privacy issues, see the *Forum Guide to Protecting the Privacy of Student Information: State and Local Education Agencies (http://nces.ed.gov/forum/pub_2004330.asp), Privacy Issues in Education Staff Records (http://nces.ed.gov/forum/pub_2000363.asp),* and the other privacy-related resources available from the National Forum on Education Statistics *(http://nces.ed.gov/forum/ferpa_links.asp).*

[3] Visit *http://www.ed.gov/policy/gen/guid/fpco/ferpa/index.html* for more information about FERPA.

[4] Visit *http://aspe.hhs.gov/admnsimp/pl104191.htm* for more information about HIPAA.

[5] Visit *http://www.access.gpo.gov/uscode/title20/chapter33_.html* for more information about IDEA.

Part IV
SUMMARY

This document addresses the following questions:

Part I. What is a Decision Support System?

- What is a decision support system?
- How does a data decision support system differ from a data warehouse and a data mart?
- What types of questions might a decision support system address when used in an education organization?

Part II. Components of a Decision Support System

- What components, features, and capabilities commonly comprise a decision support system?
- How does each broad category of these components and features contribute to system operation?

Part III. Developing a Decision Support System

- How does an education organization buy or develop a decision support system?
- How are stakeholders trained to use a decision support system?

A decision support system is a cohesive, integrated system of hardware and software designed specifically to manipulate data and enable users to support problem solving and decisionmaking by drawing useful information from disparate sources of data. Generally, decision support systems incorporate the full spectrum of information management practices and components–data collections; hardware, networks, and operating systems; underlying data sources; extract, transform, and load (ETL) processes; a data warehouse or data aggregator; analysis and reporting tools; and a user dashboard.

Implementing a decision support system is a proactive way to use data to manage, operate, and evaluate education institutions. Depending on the availability and quality of the underlying data, such a system can be used to address a wide range of issues, including academic questions about the classroom, operational questions beyond the boundaries of the classroom, and policy questions at the state level.

Purchasing or developing a decision support system is a major undertaking. Education organizations interested in such a purchase or development will need to plan and implement steps common to any major technology acquisition, including conducting a needs assessment, planning for data security, and training stakeholders.

The *Forum Guide to Decision Support Systems: A Resource for Educators* was written in response to the shortage of reliable and objective information about the use of decision support systems in education organizations. It is intended to help "educate the educators" about decision support systems.

Appendix
ELEMENTS OF A DECISION SUPPORT SYSTEM REQUEST FOR PROPOSAL (RFP)

While this document does not provide detailed information on how to write a Request for Proposal (RFP), the following types of issues may be included in a decision support system RFP. Specific requirements will, of course, vary based on each organization's unique needs and circumstances.

Scope of Work

One of the most important steps of good RFP preparation is careful and complete planning. If the organization has staff with expertise in decision support systems, these individuals might be able to identify and detail many of the design features likely to be important to potential users. Planners may even wish to generate a formal technical and functional specification document (see part III) to inform the RFP. They might also consider having representatives look at systems in peer organizations to learn about potential pitfalls as well as features and processes that have been particularly useful. If this is impractical, or if no staff members are qualified to do this research, hiring technical professionals might be advisable to help define the district's needs and write the RFP.[1]

Whether the organization is purchasing an off-the-shelf product or soliciting a custom-built decision support system, the RFP should address the following scope of work issues.

Components of Decision Support System Technology

The RFP should address the planning, development, testing, and verification of at least five major technology components.

1. **An extract, transform, and load (ETL) mechanism.** This may simply be software code that permits data from multiple formats to be brought into the data warehouse through a manual or automated process; or it may involve integration servers that allow one part of the system to communicate with another and extract information through transparent processes.
2. **A data warehouse.** This repository should allow data from multiple sources to be housed for manipulation and analysis after ETL; or it may be a data aggregation tool that directly accesses existing, interoperable databases.
3. **An analysis tool.** An application that can aggregate and disaggregate data, perform statistical analysis or other procedures, and transform raw data into useful information.

4. **A reporting tool.** An application that can translate analytical results into graphic, tabular, or other formats that permit users to more easily read or comprehend results and communicate them clearly to policymakers, parents, community members, educators, and others who may have a need to know or an interest in the result.
5. **A user dashboard.** An interface customized to meet the information needs of key user groups such as teachers, principals, administrative staff, policymakers, and parents.

Depending on vendor capabilities, multiple RFPs to multiple vendors may need to be combined to ensure optimum system performance. However, the end product must be a single, fully integrated decision support system. Information needs, staff technical expertise, current data architecture, cost, and other factors unique to the organization will dictate optimal configuration.

System Needs

- **Performance features.** What, specifically, must the decision support system do for each user group? All necessary features should be specified in the RFP, including in-and-out data transfer, statistical or other analysis capabilities, report formatting, and automated tasks.
- **Scalability.** How much computing power does the organization's present technology system use, and what is the anticipated growth for its expected life?
- **Interoperability.** What other data systems, hardware, and software will be part of the decision support system? Must data transfer be automatic and seamless, or are manual export and import acceptable for data exchange?
- **Networking.** Must the data management system operate on the Internet, on an intranet, or through another closed networking system?
- **Ongoing support.** What level of support is available, and at what cost? What are the hours of operation for these services, and how quickly will the vendor respond to system problems?

Legal Issues

- **Project completion.** How is the organization protected if the system developer fails to adequately complete contractual obligations? What happens if the developer goes out of business?
- **Ownership of code.** Does the organization have the right to modify components and share source code with other entities, as needed, inside or outside of the organization?
- **Copyright protection.** What is the organization's legal protection if the vendor violates copyright or patent laws?
- **Mediation of disputes.** What are the rules to settle any disputes that may arise during development/installation or after system implementation?

Deliverables and Cost

- **Cost basis.** Will the proposal involve a fixed cost for the entire scope of the project or a negotiated price that may be adjusted during development?
- **Schedule for deliverables and progress payments.** What is a reasonable schedule for developing a customized system, or for modifying and installing an off-the-

shelf system? Are deliverables and deliverable dates clearly and reasonably identified? What documentation will be required to trigger incremental payments for services rendered? On what schedule? In what increments?

- **Timelines and liquidated damages.** What guarantees can the vendor provide that the system will be ready to run by the required date? Should the organization consider imposing a financial penalty if the vendor fails to meet the promised schedule?
- **Total cost of ownership.** How might licensing and leasing agreements (rather than purchased components) affect the total cost of ownership? Can pricing be locked in for a reasonable period of time (e.g., three years)?
- **Documentation.** What documentation (proof of ownership, instructions for use, warrantees and guarantees, etc.) does the organization require? Will source codes and data dictionaries be included in these documentation requirements?

Professional Development

- **Initial training.** Who is responsible for training staff and non-staff users? Does the professional development include all types of users, from the superintendent and school board members to parents and community members? Will the modules use actual data from the organization for training or predefined, generic data to simulate use?
- **Ongoing training.** What provisions are in place for ongoing training? How will new users and staff members be trained? Will vendors offer additional training after system upgrades?

Note

[1] For more information about developing a technical RFP, see *Weaving a Secure Web Around Education: A Guide to Technology Standards and Security* (NCES 2003–381). National Forum on Education Statistics. (2003). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Available online at *http://nces.ed.gov/pubs2003/secureweb/ch_5.asp*.